Philosophia Scientiæ
Travaux d'histoire et de philosophie des sciences

16-1 | 2012
From Practice to Results in Logic and Mathematics

Towards a Practice-based Philosophy of Logic: Formal Languages as a Case Study

Catarina Dutilh Novaes

Electronic version
URL: http://journals.openedition.org/philosophiascientiae/719
DOI: 10.4000/philosophiascientiae.719
ISSN: 1775-4283

Publisher
Éditions Kimé

Printed version
Date of publication: 1 April 2012
Number of pages: 71-102
ISBN: 978-2-84174-581-4
ISSN: 1281-2463

Electronic reference
Catarina Dutilh Novaes, « Towards a Practice-based Philosophy of Logic: Formal Languages as a Case Study », *Philosophia Scientiæ* [Online], 16-1 | 2012, Online since 01 April 2015, connection on 30 April 2019. URL : http://journals.openedition.org/philosophiascientiae/719 ; DOI : 10.4000/philosophiascientiae.719

Towards a Practice-based Philosophy of Logic : Formal Languages as a Case Study *

Catarina Dutilh Novaes
University of Groningen (The Netherlands)

Résumé : Au cours des dernières décennies, les travaux portant sur les pratiques humaines réelles ont pris de l'importance dans différents domaines de la philosophie, sans pour autant atteindre une position dominante. À ce jour, ce type de tournant pratique n'a cependant pas encore pénétré la philosophie de la *logique.* En première partie, j'esquisse ce que serait (ou pourrait être) une philosophie de la logique centrée sur l'étude des pratiques, en insistant en particulier sur sa *pertinence* et sur la *manière* de la conduire. En deuxième partie, j'illustre cette approche centrée sur les pratiques au moyen d'une étude de cas : le rôle joué par les langages formels en logique, en particulier dans les pratiques des logiciens. Ma thèse est que les langages formels jouent un rôle *opératoire* fondamental dans le travail des logiciens en tant que technologie pratique du crayon et du papier, génératrice de processus cognitifs – et qui plus spécifiquement vient contrebalancer certains de nos schémas cognitifs « spontanés » peu adéquats à la recherche en logique (ainsi que dans d'autres domaines). Cette thèse sera appuyée sur des données empiriques venant de la recherche en psychologie du raisonnement. Avec cette analyse j'espère montrer qu'une philosophie de la logique centrée sur l'étude des pratiques peut être fructueuse, en particulier si elle est complétée par les réflexions méthodologiques nécessaires.

Abstract: In different subfields of philosophy, focus on actual human practices has been an important (albeit still somewhat non-mainstream) approach in recent decades. But so far, no such practice-based turn has yet taken place within the philosophy of *logic.* In the first part of the paper, I delineate what a practice-based philosophy of logic would (could) look like, insisting in particular on *why* it can be relevant and *how* it is to be undertaken. In the

*. This research was supported by a VENI grant from the Netherlands Organization for Scientific Research (NWO). I would also like to thank two anonymous reviewers for detailed comments, and the audience at the 'From Practices to Results' conference in Nancy (June 2010) for valuable feedback.

Philosophia Scientiæ, 16 (1), 2012, 71–102.

second part, I illustrate the proposed practice-based approach by means of a case-study: the role played by formal languages in logic, in particular in the practices of logicians. I argue that formal languages play a fundamental *operative* role in the work of logicians, as a paper-and-pencil, hands-on technology triggering certain cognitive processes—more specifically, countering some of our more 'spontaneous' cognitive patterns which are not particularly suitable for research in logic (as well as in other fields). I substantiate these claims with empirical data from research in the psychology of reasoning. With this analysis, I hope to show that a practice-based philosophy of logic can be a fruitful enterprise, in particular if accompanied by much-needed methodological reflection.

In different subfields of philosophy, focus on actual human practices has been an important (albeit still somewhat non-mainstream) approach for some time already. This is true in particular of general philosophy of science, and to a lesser extent of philosophy of mathematics: following the revolutionary work of authors such as Kuhn and Lakatos, quite a few philosophers of science and mathematics sought to incorporate the actual practices of scientists and mathematicians into their philosophical analyses. Now, it seems fair to say that such a practice-based turn in these fields has delivered important results; we now have a more thorough and encompassing understanding of science as an enterprise thanks to these analyses.

In this sense, it is perhaps surprising to notice that no such practice-based turn has yet taken place within the philosophy of *logic*. Why is that? Is there something about logic that makes it intrinsically unsuitable for a practice-based approach? What are the prospects for new insight into traditional philosophical questions pertaining to logic (e.g. the nature and scope of logic) to be gained from focus on the practices of logicians? It might be contended that, given logic's normative (as opposed to descriptive) nature, and given that it deals with a priori issues, not empirical ones, attention to the practices of logicians is a wrongheaded approach from the start. The main goal of the present contribution is to argue that this is not the case, independently of the issue of the normativity and a prioricity of logic. [1]

In the first part of the paper, I delineate what a practice-based philosophy of logic would (could) look like, insisting in particular on *why* it can be relevant and *how* it is to be articulated. The 'how?' question is especially significant, as crucial methodological challenges must be addressed for the formulation of a methodologically robust practice-based philosophy of logic. In the second part, I illustrate how a practice-based approach in (the philosophy of) logic could be developed by means of a case-study: the use of formal languages in logic, in

1. In other words, attention to the practices of logicians may be illuminating even within a normative, a priori conception of logic, as I hope will become clear in the paper.

particular in the practices of logicians. (The material presented here is essentially a condensed version of the main arguments presented in book-length in [Dutilh Novaes 2012].) I argue that formal languages play a fundamental *operative* role in the work of logicians, as a paper-and-pencil, hands-on technology triggering certain cognitive mechanisms—more specifically, countering some of our more 'natural' cognitive mechanisms which are not particularly suitable for research in logic (as well as in other fields). I substantiate these claims mostly with empirical data from research on the psychology of reasoning. By means of this analysis, I hope to show that a practice-based philosophy of logic can be a fruitful enterprise, especially if accompanied by much-needed methodological reflection.

1 Practice-based philosophy of logic

1.1 Why?

Perhaps the first question to be asked before we embark on a new enterprise such as a practice-based approach in philosophy of logic is: Why do we need a new approach at all? In what sense do more traditional approaches fail to clarify certain aspects of the target-phenomenon, and in which ways is the proposed new approach likely to shed new light on the matter? In order to address these questions, let me first offer some general considerations on the very idea of philosophical analyses with emphasis on (human) practices.

Let us start with the schema 'practice-based philosophy of X', where 'X' is a given (scientific) field. If we replace 'X' with 'science', we obtain 'practice-based philosophy of science', which was arguably the first practice-based approach systematically developed within philosophy. Generally speaking, a practice-based philosophy of science takes actual scientific practices as a starting point instead of an idealized, abstract notion of 'science'. The contrast between Popper's and Kuhn's respective views is illustrative: Popper maintained that science proceeds (indeed, *should* proceed) by means of falsification [Popper 1959], but Kuhn pointed out that this picture was in no way a reflection of *actual* scientific practices [Kuhn 1962]. A Popperian might still contend that this is in any case how science *should* proceed; but when there is such a significant mismatch between theory and object of analysis, in what sense is the theory still informative or relevant in any interesting way?

The crucial gestalt shift undertaken by Kuhn was arguably that of viewing 'science' not as an abstract entity—the collection of scientific theories—but rather as a body of (rather heterogenic) *human practices.* From a Kuhnian perspective, such practices are the actual phenomena that a philosopher of science should address, not (only) the theories themselves and their logical relations as such. But in order to argue that the Popperian picture was a wrong picture of scientific practice, Kuhn needed to make his philosophical analyses *empirically-informed*: it was on the basis of investigations on the *history of science* that Kuhn could challenge the Popperian picture and argue that, rather than falsification, a scientist typically aims at the *confirmation* of theories. So here is the first important feature of (and challenge for) a practice-based approach in philosophy: it must be empirically informed, as it must gather data on how e.g. scientific research is *actually* conducted. This may seem to jeopardize the very core of philosophical methodology as traditionally construed, and to blur the distinction between philosophy, sociology and history. But are these insurmountable difficulties?

Besides practice-based philosophy of science (which also has Hacking and Latour, among others, as prominent names), recent years have witnessed the emergence of a practice-based philosophy of mathematics in the work of e.g. J. P. van Bendegem [Kerkhove & Bendegem 2007], Hersh [Hersh 1997], [Hersh 2006], Mancosu [Mancosu 2008], among others. Again, the main idea is that traditional philosophy of mathematics, with the predominance of the 'foundations of mathematics' research program, offers a distorted or in any case extremely limited picture of mathematics as actually practiced by mathematicians. For philosophers interested in mathematical practices, the target-phenomenon to be accounted for is not only the mathematicians' *output* (the theories they formulate and the theorems they prove), but also the *processes* (both individual and social) leading to the output, i.e. the practices of mathematicians.

In a sense, in both cases the crucial step seems to be a widening of the scope of philosophy of X so as to include the actual practices of the practitioners of X. It may be debated whether this new scope is meant to *replace* the old, output-oriented scope, or whether it is meant to be an *addition,* not necessarily seeking to supplant traditional forms of philosophical investigation of X. But ultimately, practice-based approaches to science and mathematics seem to rely crucially on an *assumption* and on an *empirical claim*: actual practices of mathematicians and scientists are philosophically relevant (assumption), and traditional philosophical accounts of both science and mathematics (those with no particular focus on practices) do not reflect actual practices within these fields (empirical claim).[2] I believe the empirical claim to have been essentially vindicated by the works of Kuhn and many others following his footsteps (even though there is of course still scope for debate on the exact details of encompassing models of scientific and mathematical practices). As for the assumption, one way that it could be discharged is if a practice-based approach to X succeeds in solving, or at least in shed-

2. Notice that the distinction between process vs. output does not correspond neatly to the 'context of discovery vs. context of justification' distinction, as one of the elements that is of interest from a practice-based perspective is how canons of scientific justification are established and enforced by communities of practitioners.

ding new light on, an open question within traditional philosophy of X, one with respect to which traditional philosophical methodology has proved to be insufficiently illuminating. [3]

Let us take a closer look at the current state in the philosophy of logic. I think it is fair to say that 'traditional' philosophy of logic is often far removed from the actual, latest developments of research in logic properly speaking. The main focus is still on traditional themes such as truth, consequence, paradoxes, etc. To be sure, such topics remain of crucial importance, but there is a wide range of developments within logic as a discipline that receive scarce attention from philosophers of logic. Indeed, the logic that they talk about is all too often the logic of several decades ago, when (mathematical) logic was almost exclusively concerned with the foundations of mathematics. In a sense, philosophy of logic is still very much tied to what could be described as a 'Quinean agenda', and in many cases still to the idea that first-order predicate logic is the quintessential logical system (as in e.g. [Sher 2008]). In practice, however, and in particular in terms of applications, first-order logic is much less prominent, a fact perhaps related to its sub-optimal computational properties. In short, current *philosophy* of logic does not pay sufficient attention to the actual, current practices of logicians.

But what are the current practices of logicians? Again, this is an empirical question, and one which is not likely to receive an unproblematic answer, precisely because what is to count as 'logic' is a contentious issue to start with. Delineating what is to count as a 'practice of logic', even if not in a sharp, clear-cut way, is most certainly not a trivial enterprise. While I acknowledge this difficulty, I think a few general remarks can at least convey some of the mismatches between 'logic' as talked about by philosophers of logic and 'logic' as practiced by logicians. Firstly, logic is no longer exclusively concerned with the foundations of mathematics; it intersects with areas such as computer science, game theory, decision theory, linguistics, cognitive science, etc. Secondly, several different logical systems besides first-order logic are regularly used and studied, but discussions on logical pluralism do not seem to be able to really make sense of the plurality of logical systems. Arguably, explaining how (if at all) the current plurality of logical systems available in the market can (or cannot) be harmonized with the traditional idea of 'logic as the science of correct reasoning' is one of the most pressing tasks for philosophers of logic. While there has been quite some interest in this issue on the part of philosophers [Haack 2000], [Beall & Restall 2006], a hallmark of such discussions on logical pluralism is that they are arguably still not sufficiently

3. There are other ways of discharging the assumption, such as raising new, important questions which are not (cannot) be formulated within traditional philosophy of X (I owe this point to an anonymous referee). But in such cases, traditional philosopher of X can still argue that the question is not in fact relevant, or that it is not even a truly *philosophical* question (speaking from personal experience!).

'pluralistic', leaving too much of the actual practices of logicians out (as argued by e.g. [Benthem 2008]). Typically, only those that could be described as the 'classical' non-classical logics are explicitly discussed (intuitionistic logic, relevant logic, and in recent years paraconsistent logics). In particular, actual research in logic goes well beyond the concepts of consequence and truth, but nevertheless discussions on logical pluralism are typically formulated in terms of consequence-pluralism and alethic pluralism.

I am not suggesting that current discussions within philosophy of logic are no longer relevant and should simply be abandoned in favor of a practice-based approach. Rather, I am suggesting that the mismatch between philosophical theorizing and actual (recent) practices when it comes to logic is a reason for us to take seriously the idea of expanding the research agenda within philosophy of logic, and in particular to pay more attention to the actual practices of logicians. The idea is to raise philosophical issues pertaining to the latest developments in logic,[4] and to ask ourselves questions such as: What are the cognitive and social mechanisms involved in the development of logic as a discipline? What is it that we do when we 'do logic', and how do we do it?

So here is a tentative characterization of practice-based philosophy of logic in its goals and tasks:

- It takes as its starting point logic as it was and is actually practiced—recent developments as well as its history.
- Its tasks are:
  - to clarify underlying assumptions, raise pertinent questions;
  - to draw philosophically significant conclusions from technical results,
  - to 'make sense' of logical practices: the tools and technologies involved, the cognitive processes, social interactions.

Let me stress once again that practice-based philosophy of logic need not replace more traditional forms of investigation in the field; it is best seen as an alternative, complementary philosophical approach to logic, but one which may ultimately even shed new light on traditional issues.

1.2 How?

Now that I have addressed the 'why' question, let us turn to the much more delicate 'how' question: how is practice-based philosophy of logic to be developed in a methodologically robust manner? Mere anecdotal evidence will obviously not be sufficient; but as already noticed, a practice-based philosophy of logic *must* be empirically-informed, given that the goal is precisely to

4. One field that must be further explored on philosophical grounds is the connection between logic and games. There are of course several logical systems that make heavy use of the game analogy (dialogical logic, game-theoretic semantics, etc.), but is the analogy justified? In which way is logic really (like) a game? See for example [Marion 2009] and [Hodges 2009] for discussions of these issues.

take actual practices into account. [5] Here again we encounter the output vs. process distinction: while traditional philosophy of logic focuses on the output of logical practices, i.e. the logical theories themselves, from a practice-based perspective logic is viewed as a body of human cognitive and social processes *as well as* a collection of theories. From this point of view, how logicians arrive at their results—the heuristics of logical practices—becomes an important question, and insofar as they are human beings like any other, the general cognitive makeup of human agents and their patterns of social interaction will be significant.

Indeed, I submit that there are essentially two intertwined but distinct levels of practices to be investigated: the social level and the individual level. On the one hand, the social level of logic as a collective, public enterprise involves networks of people who communicate with each other and whose work builds on previous work (it is a cumulative enterprise). Naturally, logicians share specific (socially established) norms on how work in logic ought to be done. On the other hand, the individual level of logic as a cognitive enterprise regards the cognitive processes taking place when a logician is at work; even though the social aspect is of course fundamental for the creative process, ultimately thinking remains an individual, private matter. I suggest that these two levels may require distinct methodologies to be investigated.

With respect to the social level, a natural question to be asked is: what distinguishes a practice-based philosophy of logic from a sociology of logic? Granted, sociology of logic is an almost non-existing field, but at least one comprehensive sociological study of the practices of logicians has been undertaken with significant results, namely [Rosental 2008]. Would a practice-based philosophy of logic focusing on the social level be significantly different from such a sociological approach? I believe it would, even though a philosophical analysis would do well to be *informed* by the results obtained by means of the application of solid sociological methodology such as Rosental's. But just to illustrate the kind of analysis that a practice-based philosophy of logic could offer on the social level, let us reflect for a moment on the ubiquitous practice of presenting a proof-system and a semantics, and then moving on to proving soundness and completeness results. That this is how we 'do logic' is such a widespread convention that few people reflect on what exactly it means to offer a soundness and completeness proof from a philosophical point of view. A practice-based philosopher of logic could raise precisely this kind of question: why is this manner of proceeding so widespread, and what exactly is achieved by means of such proofs? Of course, there are traditional philosophical analyses of the significance of proofs of soundness and completeness, such as [Dummett 1978], but they fail to highlight the fact that it has to a

5. Insofar as logic is traditionally seen as an a priori, normative enterprise, it might be thought that 'empirically-informed philosophy of logic' is something of an oxymoron. To argue that it is not so is one of the goals of the present contribution.

large extent become a *social* norm to proceed in this manner. A practice-based philosopher of logic, by contrast, may outline the partially conventional nature of such a procedure, and may even go as far as questioning the propriety of this procedure in specific cases.[6]

In short, the main difference between a sociology of logic and a practice-based philosophy of logic is that the former is fundamentally *descriptive*; the latter, while having a descriptive dimension (and possibly being informed by sociological findings), will also be to some extent *prescriptive.* Some of the tasks of a practice-based philosophy of logic are to offer *critical* analyses of the conceptual foundations of actual work being done in logic (clarifying underlying assumptions), and to identify possible conceptual problems underlying the practices and to suggest directions for improvement. Indeed, actual practices are not always (necessarily) 'right'.

What kind of methodology can we use for data-gathering on the practices of logicians? As already mentioned, mere anecdotal evidence is not sufficient. Moreover, once such data are gathered, it is not entirely obvious that we can then proceed with the 'usual' philosophical methodology, whatever that may be. At this point, however, it would be premature to offer full-fledged methodological considerations; faithful to the practice-based principle, I believe that it is only by *doing* practice-based philosophy of logic that the appropriate methodological guidelines will really present themselves.

Nevertheless, even at this early stage it is important to be clear on the kind of methodology that can be used for data-gathering on practices. I have a fairly precise idea on how to get started on the individual level of logic as a cognitive enterprise: a promising approach seems to be to take into account findings from cognitive science and psychology on how humans *in general* reason. There is a well-developed tradition in the psychology of reasoning which has gathered significant data on human reasoning patterns, and an interesting question is precisely to inquire into the cognitive changes produced by logical training on this initial cognitive state that logicians, as humans, share with all other humans (it is precisely the line of investigation that I shall pursue in the second part of the paper). Although little work has been done on the psychology and cognitive science of learning logic specifically, there is *some* work on the topic; [Stenning 2002], for example, is a remarkable illustration of how cognitive processes engaged in the learning of logic can be studied.[7] Moreover, the

6. I have in mind the cases in which a semantics does not naturally present itself and a somewhat ad hoc, concocted semantics is developed to fill in the gap, even though this semantics does not really represent the original target-phenomenon faithfully. See [Andrade-Lotero & Dutilh-Novaes 2010] for a critical discussion of the significance of proofs of soundness and completeness in the case of Aristotelian syllogistic.

7. There is also an interesting tradition in research on mathematics education, on how for example the concept of logical necessity develops upon instruction; see for instance [Morris & Sloutsky 1998].

practice-based philosopher of logic already has a good starting point with the available data on how humans in general, rather than logicians specifically, reason. I shall argue in the second part of the paper that significant insight can be gained on the practices of logicians just by taking this material as a starting point.

As for the collective level of logic as a social enterprise, data-gathering seems a delicate matter. Robust sociological methodology would have to be employed, such as e.g. systematic, quantitative analyses of the corpus of articles published on a given topic in view of a specific question to be addressed. Historical studies may also prove to be useful, as they may provide information on the variants and invariants in the practices of logicians across different times and places. But at this point, data-gathering for the social level of analysis would require significantly more discussion than I can offer here, and I hope that others may be able to offer suggestions for further development in this area.

In general, the methodological challenges for the formulation of a practice-based philosophy of logic stem from the fact that the approach finds itself in the difficult position of balancing descriptive and prescriptive elements. It must not only deal with the is/ought divide, but it must also find a way to merge the 'is' level with the 'ought' level. It does not take a purely idealized notion of what logic ought to be as its starting point, but it is not meant to be sociology of logic either. Perhaps one way to describe this predicament is the following: practice-based philosophy of logic discusses how things *ought to be* (with respect to logical practices), but taking as a starting point how things *actually are*—a constrained, situated form of normativity. Moreover, ideally the approach should be able to create a situation of *reflective equilibrium* between practices and theory: ideally, the dialogue between the practice-based philosopher of logic and the logician should go both ways.[8]

2 A case study: uses of formal languages in logic

2.1 Why do logicians use formal languages?

One does not need to be a practice-based logician to recognize the importance of understanding the role of formal languages in logic. True enough, most of us seem to take for granted that 'this is simply how we do things', forgetting (or dismissing) the fact that, for most of its history, logic was not practiced with formal languages as we now know them. Yet, this seems to be a fundamental

8. Admittedly, this might be too much to hope for; it is not obvious that the logician will be interested in talking to the practice-based philosopher of logic.

question: why use formal languages when doing logic? What is the actual impact of using formal languages for research in logic? Does it really make a difference for the logician's investigations in terms of the results obtained? Is it necessary for the development of logic as a discipline? One cannot deny the substantial changes that logical practices underwent since it became customary to do logic with formal languages, and this phenomenon requires an explanation. Although perhaps not indispensable for logical practices up to a certain level of abstraction, formal languages are arguably such powerful tools that they profoundly transform the ways in which we do logic.

Typical answers to these questions refer to the clarity of expression and level of abstraction afforded by formal languages; sometimes reference is also made to the object-language vs. meta-language distinction, which greatly contributed to the development of meta-theoretical investigations. It seems to me however that, while such remarks are not false properly speaking, in fact they only 'scratch the surface' of the phenomena in question.[9] There seems to be much more to the use of formal languages in logical practices than these elements alone, and in particular we should inquire into the *cognitive* impact of using formal languages when doing logic.

First, here is a question that is often seen as either trivial or pedantic (or both): what are formal languages? Some may dismiss the question as unimportant (e.g. [Avron 1994, 218]), while others may say that there is an exceedingly simple and philosophically uninteresting answer. Indeed, from a purely mathematical point of view, a formal language is characterized by a finite vocabulary and a finite and sharply defined syntax: its (well-formed) expressions and formulas are generated recursively on the basis of the vocabulary and the syntax (forming a potentially infinite class). Again, this is of course true, but it leaves a lot unexplained. For example, in what sense are formal languages *formal,* and in what sense are they *languages*? Do they fulfill a communicative role as vernacular languages do? Do they have a social dimension as languages used for communication among logicians? Presumably, the adjective 'formal' differentiates them from other kinds of linguistic systems.[10] Elsewhere [Dutilh Novaes 2011], I discuss at length the different meanings that the term 'formal' has with respect to logic, and it would seem that, of the eight senses of 'formal' identified in that paper, formal languages are formal in particular in the sense of formal as *de-semantification* (a concept to be explained below) and in the sense of formal as *computable* (related both to how the syntax of a formal language operates and how inference-making with a formal language and an appropriate deductive system typically functions). These two aspects will have a key impact on the cognitive effect of doing logic with formal languages, as I shall argue.

9. I argue for these claims in more detail in [Dutilh Novaes 2012, chap. 3].
10. Although Montague famously claimed that English is a formal language.

But before focusing on these cognitive aspects, a few more preliminary observations on formal languages and logical practices seem in place. First of all, formal languages obviously do not replace other forms of communication in logical practices: typically, logicians use a mix of formal and vernacular languages, switching a bit back and forth when convenient (this can be observed in particular in oral contexts). But again, research in logic which uses formal languages is significantly different from research in logic without them (as history shows); the appearance of this specific technology has drastically changed the way logicians work, not only in terms of methods but also in terms of the results obtained.

But what kind of tool is it? Some fairly obvious but often overlooked facts about formal languages are worth mentioning here. To begin with, formal languages are *written* languages, with no obvious spoken counterparts. As such, they involve predominantly (but perhaps not fundamentally) our *visual* capacities. Equally important is the observation that formal languages came into existence only after a very long and gradual process going through the use of schematic letters and the development of mathematical notation (algebra in particular), spanning over many centuries and different continents [Staal 2006], [Dutilh Novaes 2012, chap. 3]. The 17$^{\text{th}}$ century is a crucial period in this sense, not only for the groundbreaking developments in notations for algebra (in particular with Viète and Descartes), but also in that, perhaps for the first time in history, logic and mathematics were increasingly viewed as related disciplines [Mugnai 2010]. This greatly facilitated the transfer of mathematical notational techniques to logic. And yet, as is well known, the development of the first full-fledged logical formal language had to wait until 1879, with Frege's *Begriffsschrift.*

But the *Begriffsschrift* did not have much of an impact on logical and mathematical practices in first instance. This sociological fact may be related to the stage of development of the foundational program at the time; in 1879, axiomatizations of different subfields of mathematics were in their very early days, and at that point the focus was on the axioms rather than on the rules of inference underlying the theories (which were then informally and tacitly adopted, as if they were unproblematic [Awodey & Reck 2002]). In other words, at that point, the need for such a formal language as Frege's concept-script was simply not truly felt. But by the time the first volume of Whitehead and Russell's *Principia Mathematica* appeared (which the authors themselves saw as a continuation of Frege's project), the time was right for a more 'formal' approach to rules of inference, in particular because meta-theoretical questions became increasingly seen as fundamental.[11] More specifically, it became clear that the completeness (in different senses of 'completeness'—see [Awodey &

11. This does not mean that *Principia Mathematica* itself had already reached the stage of formulating meta-theoretical questions (essentially, it had not); rather, the claim is that the meta-theoretical questions which were beginning to emerge (as

Reck 2002]) of the proposed axiomatizations could not be taken for granted and had to be established by means of proofs. Now, to conduct such meta-theoretical investigations on those axiomatizations, it became important to 'objectify' the underlying logic as a mathematical object itself, and thus the concept of object-language emerged.

The classical rationale for the introduction of formal languages is the idea that vernacular languages are prone to all kinds of ambiguities and imperfections, and a formal, artificial language would be the result of a sanitation of expressive means for the purposes of scientific investigation in particular. The preface of Frege's *Begriffsschrift* is the *locus classicus* for this position (I shall say more on the expressive role of formal languages shortly). Moreover, a crucial and essentially new function for formal languages was discovered in the early 20$^{\text{th}}$ century: to facilitate meta-theoretical investigations to be conducted.[12] While these two aspects are undoubtedly significant to explain the impact of the use of formal languages in logical practices, my main contention here is that they only tell part of the story, and my goal is to propose at least a partial account of the rest of the story.

2.2 Formal languages as a technology

Here is yet another hypothesis which may not sound particularly illuminating at first, but which I believe truly contributes to clarifying the role of formal languages in logical practices: formal languages are best seen as a particular kind of *technology.*

Of course, this hypothesis will not be very informative unless we can provide a more precise meaning to the rather vague term 'technology'. As a first approximation, a technology can be described as a specific method, material or device used to solve practical problems. Formal languages as such are not a method by themselves, but they are devices that allow for the implementation of certain methods. Ultimately, formal languages are a valuable item in a logician's toolbox; but naturally, being tools, they will deliver results only if put to use in fruitful ways.

A particularly puzzling aspect of technologies in general is that they are usually developed with specific applications in view (i.e. to fulfill certain needs that the hitherto available technologies fail to address), but often turn out to offer possibilities that had not been originally foreseen, and which go beyond the specific practical problems they were created to address. We could

described in [Awodey & Reck 2002]) could be more easily addressed by others against the background offered by *Principia Mathematica.*

12. But notice that formal languages are not a necessary condition for meta-logical investigations. The *Prior Analytics* itself, the first logical text in this tradition, offers sophisticated meta-theoretical analyses.

refer to this phenomenon as the 'surprise factor' in any technological development. What this means is that, once a given technology is developed to tackle a specific familiar issue, it often opens up a whole new range of possibilities which would otherwise not even be conceivable prior to its development. New technologies may literally create new worlds, besides bringing answers to existing problems. [13]

Perhaps the best example of a radically life-changing technology is writing. Writing developed through thousands of years, but for a very long time, it consisted in very rudimentary forms of proto-writing [Schamdt-Besserat 1996]. Originally, proto-writing was developed in Mesopotamia as a simple technology for accounting, i.e. for keeping track of production surplus and goods in general; naturally, such needs only arose against the background of a society which had abandoned a hunting-gathering economy in favor of agriculture and herding. [14] So what started as rudimentary techniques to keep track of goods slowly but surely developed into a much more powerful technology, one which profoundly modified the way human beings (in literate societies) live their lives.

Formal languages are, as already mentioned, *written languages,* so the history of their development is a specific chapter in the general history of the development of writing. An important aspect is how the development of mathematical notation in particular went hand in hand with the search for notations that would facilitate processes of *calculation*; this is especially conspicuous in the development of numerical notation—see [De Cruz & De Smedt 2010]. Interestingly, with respect to numerical notation, expressivity and ease of calculation often do *not* go hand in hand. The classical example is the contrast between Roman numerals and Indo-Arabic numerals; the former are in a sense more intuitive and 'iconic', but they constitute a cumbersome tool for calculation, while the opposite is true of the latter. [15] Indeed, there seems to be an inherent tension between desiderata of expressivity and desiderata of effective calculation, a tension which will occupy a central position in the present analysis.

As already noticed, the pioneers of the introduction of formal languages in logic typically had concerns of expressivity (precision, objectivity) in mind when defending the need for a technology other than vernacular (written) languages in their investigations. The notable exception is of course Leibniz, for whom the calculative aspect of an artificial language such as his envisioned

13. Needless to say, the surprise factor in technological development can also lead to disastrous consequences.

14. Writing did arise independently at a few other places and times, but the history of such developments is not significantly different (as far as we can establish from our current knowledge of these developments).

15. See [Krämer 2003, 531]. But compare [De Cruz, Neth & Schlimm 2010] on the cognitive impact of different numerical systems.

lingua characteristic and its potential to lead to new discoveries was crucial. The claim I will be defending in what follows is that this facilitating effect for logical reasoning as such is the somewhat 'unexpected' (but not for Leibniz, of course) upshot of the introduction of the specific technology corresponding to formal languages in logical practices. Inquiring into the cognitive processes underlying these phenomena is one of the purposes of the paper. But rather than using the term 'calculative', whose meaning is more restricted than what I have in mind, I will be using the term 'operative' to refer to this feature of formal languages.

In effect, we can discern three different and somewhat conflicting roles of formal languages in the practices of logicians (but notice that this enumeration does not make any claim of exhaustivity): [16]

– Expressive role
– Iconic role
– Operative (calculative) role.

The expressive role of formal languages. A discussion of the expressive aspect of formal languages must begin with an examination of Frege's own description of the purpose(s) of his 'ideography':

> My first step was to attempt to reduce the concept of ordering in a sequence to that of logical consequence, so as to proceed from there to the concept of number. To prevent anything intuitive from penetrating here unnoticed, I had to bend every effort to keep the chain of inferences free of gaps. In attempting to comply with this requirement in the strictest possible way I found the inadequacy of language to be an obstacle; no matter how unwieldy the expressions I was ready to accept, I was less and less able, as the relations became more and more complex, to attain the precision that my purpose required. This deficiency led me to the idea of the present ideography. Its first purpose, therefore, is to provide us with the most reliable test of the validity of a chain of inferences and to point out every presupposition that tries to sneak in unnoticed. [Frege 1977, 5–6]

The imperfections of vernacular languages (which Frege refers to as '*Sprache des Lebens*', the language of life) include, but are not limited to, the phenomena of ambiguity, equivocation, empty names, etc. For a variety of reasons, the expressivity afforded by vernacular languages is inadequate for Frege's project of providing logical foundations for mathematics, and in particular to test the correctness of already existing mathematical theorems on the basis of an examination of the logical validity of the corresponding chains of inference. A crucial aspect of this enterprise is the requirement of *perspicuousness*: every single step must be made absolutely explicit, as shortcuts are not allowed,

16. These three roles are discussed in more detail in [Dutilh Novaes 2012, chap. 3].

and hidden presuppositions are not welcome. It is in this sense that this motivation for using formal languages can be described as related to concerns of *expressivity*: every small inferential step and every premise/assumption must be explicitly expressed.

While I am prepared to grant that the use of a formal language can greatly facilitate the realization of these goals, it is crucial to observe that a formal language is not a *necessary* prerequisite for the success of the enterprise (and clearly, it is not sufficient either). In theory, and perhaps even in practice, this could be accomplished by means of a *regimentation* of the vernacular language in question. (Indeed, in the later Latin medieval tradition, the Latin being used for logical investigations was no longer a form of vernacular but rather a highly regimented version of Latin, containing a wide range of conventions introduced in order to increase precision of expression.)

Moreover, it is worth noticing that replacing a vernacular language with a formal one for whatever purposes typically does not come for free; there is always the risk of overall *loss* of expressivity if the formal language fails to perform its expressive function adequately. Frege himself notices that a formal language is like a microscope, i.e. a tool suitable for particular applications and in particular circumstances: but if the microscope is malfunctioning, one would be well advised to stick to the not entirely adequate, but in any case widely tested and reliable alternative, namely the eye. For these and other reasons, I believe that the gain in clarity that is (purportedly) associated with the use of formal languages (in logic and elsewhere) is by no means the main reason why using or not using formal languages has such an impact on the practices of logicians and on the results they obtain.

The iconic role of formal languages. [17] Formal languages are typically sentential languages, which means, according to Stenning's [Stenning 2002, 51] useful distinction, that their very appearance on paper as concatenations (i.e. linearly presented) must be interpreted indirectly, i.e. on the basis of the syntactical conventions which define how the inscriptions on paper must be 'scanned' in the first place. Diagrams, by contrast, rely on patterns of spatial relations between inscriptions other than concatenation, and these spatial relations can be interpreted directly as (somehow) mirroring the spatial relations present in the phenomenon being depicted.

Nevertheless, as Stenning himself acknowledges, the distinction formulated in these terms is not necessarily sharp, and many systems of representation have elements of both forms of representation, sentential and diagrammatic. 'Vernacular writing' is (normally) linear and one-dimensional (either vertically or horizontally), a fact that tends to obscure its iconic nature, as argued by

17. Given that the iconic role of formal languages is not the main topic of the present contribution, my discussion here will be rather succinct. The brief considerations presented here should essentially be seen as suggestions for future work.

Krämer [Krämer 2003]. Every form of writing is inexorably iconic insofar as presupposes inscriptions on surfaces, but given that most of them are fundamentally based on concatenation (linearity), it is all too easy to disregard the iconic element.[18] Now, what is often the case with formal languages is that, even though they are still essentially sentential languages, they nevertheless explore the iconic feature of written languages more deeply than vernacular written languages. 'Formal writing', when at its best, makes full use of the two dimensional possibilities of a surface; for example, proofs are typically represented by two-dimensional structures such as trees and graphs. Frege's *Begriffsschrift* is a good example of a notation that incorporates iconicity into a sentential form of representation (admittedly, it is a type-setting nightmare, which might explain why it was never widely adopted). As recently argued by D. Macbeth (this volume), the diagrammatic aspect of Frege's notation is fundamentally integrated into the very machinery of the system. In a similar vein, diagrammatic logical systems have been extensively studied in recent years [Shin 2002], [Shin & Lemon 2008].

The visual-iconic nature of formal logic suggests cognitive connections between doing logic and our visual faculties, and these have been explored in detail by K. Stenning in *Seeing Reason* (primarily, but not exclusively, in the context of logic teaching for undergraduate students). The cognitive impact of the use of this concrete technology—formal languages—when reasoning in logic, in particular in connection with its inherent visual and iconic nature, is a fascinating topic that deserves to be explored much more widely. Stenning's pioneering work has brought to the fore several aspects of this interaction, but it has also shown that our understanding of the impact of using different systems of representation to reason in and about logic is still at its early stages. Particularly surprising are Stenning's findings concerning individual differences in reasoning styles and the significance of using different modalities (sentential or diagrammatic systems of representation) for the resolution of the same classes of problems [Stenning 2002, chap. 02]. Whatever it may turn out to be, the story of the impact of using different technologies when doing logic will have to incorporate individual differences, and this suggests that, on a cognitive level, there is no universal, uniform way of 'doing logic'.

The operative role of formal languages. Another dimension of the use of formal languages in logic, one which is related to, but nevertheless distinct from the iconic dimension, is what we could describe as their 'operative' di-

18. Naturally, there are different kinds (degrees?) of iconicity. The iconicity of, for example, the notation in *Principia Mathematica* is fundamentally different from the iconicity in, for example, Euclidean diagrams. Krämer's main point, which I endorse here, is that every form of writing explores and relies on graphic/iconic components. One of her examples is the use of footnotes in written texts, which clearly exploits the graphic two-dimensional possibilities of surfaces (as opposed to the one-dimensionality of speech).

mension.[19] 'Operative writing' is a concept introduced by S. Krämer [Krämer 2003], which in the case of formal languages concerns the 'paper-and-pencil' dimension which seems to play an important role in how logicians reason and arrive at new results (Krämer explicitly mentions the formal languages of logicians as examples of operative writing). She presents operative writing in the following terms:

> a medium for representing a realm of cognitive phenomena [...] a tool for operating hands-on with these phenomena in order to solve problems or to prove theories pertaining to this cognitive realm. [Krämer 2003, 522]

The conception that Krämer is criticizing as too limited is the 'phonetic' conception of writing, one which reduces writing to the role of transcription of spoken language to visual media. Clearly, formal languages do not fit into this picture, along with other forms of what she describes as 'operative writing'.[20]

One possible way of understanding the divide between operative vs. non-operative writing would concern the prior availability of the cognitive phenomena that a given portion of writing is (or is not) a report of: non-operative writing would express already available thoughts, while operative writing allows for new thoughts and ideas to come into existence, precisely through the process of *operating* with, and on, portions of writing. Prima facie, this seems plausible, but from this point of view, *any* form of writing can count as operative writing at least in some circumstances. We all know that, when writing a paper in plain vernacular, doing the actual writing is often an inte-

19. That they are distinct dimensions emerges from the observation that a given system of representation can have a high level of iconicity and yet not be suitable to be 'operated on', i.e. to facilitate and aid mental processes (again, the differences between Roman and Indian/Arabic numerals). It might be thought that iconicity should enhance reasoning performance in the sense that, given that it allows for direct interpretation (in Stenning's sense), the cognitive step of 'scanning' the inscriptions and 'imposing an abstract syntactic structure on concatenations' [Stenning 2002, 51] can be 'skipped'. In practice, however, diagrammatic systems of representation allow for a lesser degree of abstraction, which in turn affects expressibility and also calculability, as I shall argue. Moreover, at least two other factors must be taken into account: individual differences in reasoning styles, and the kind of task at hand (both of which are discussed in [Stenning 2002]). The heart of the matter is in any case that there are good reasons to distinguish the iconic from the operative dimension of formal languages.

20. "In other words, this [phonetic] definition excludes so-called 'formal languages' that construct graphical systems *sui generis* and which are all, at best, verbalized retroactively and verbalized only in a limited, fragmentary form. [...] Calculus is the incarnation of operative writing." [Krämer 2003, 522]. Tellingly, Krämer started her research career as a historian of 17$^{\text{th}}$ century logic and mathematics, in particular the concepts of calculus and calculation.

gral and necessary part of the creative process.[21] In effect, it is a different, specific characteristic of Krämer's notion of operative writing that is particularly relevant for the present attempt to grasp the cognitive impact of using formal languages when doing logic: what she describes as the process of *desemantification,* which we shall examine now.

2.3 The operative role of formal languages

Before discussing the theory behind the notion of formal languages as operative writing, let me offer a few platitudes on the practices of logicians. Firstly, 99% of logicians have a black/white board in their offices; it is rare to encounter a logician (or a mathematician, for that matter) for whom writing/scribbling frantically is not a part of his/her *modus operandi.* Indeed, writing down symbols typically plays an important role in how a logician organizes his/her thoughts and comes to new ideas and insights. In itself, this is not particularly controversial: for almost all kinds of problems or reasoning tasks, it helps a great deal if we have paper and pencil at our disposal to scribble with. Thinking and problem-solving in general are not exclusively inner mental processes; by making use of external devices which help making cognitive phenomena more concrete, we are often in a much better position to solve the tasks in question. This phenomenon has been extensively analyzed in the literature on the concepts of extended mind and extended cognition [Clark & Chalmers 1998], [Clark 2008], [Menary 2010], and writing is a prototypical example of a technology extending cognition beyond the limits of skull and skin.

Still, it does look like formal languages play a particularly significant role as a hands-on tool for discovery in logic, beyond the general reasoning-enhancing properties of writing and drawing in general. They not only play a justificatory role in contributing for the perspicuousness of the correctness of a proof, but they also perform the heuristic function of contributing to the discovery of new results and theorems, more than just any other arbitrary form of writing/drawing. Of course, we must always bear in mind that individual differences and different styles of reasoning (again, see [Stenning 2002]) suggest that this heuristic dimension does not operate uniformly on all agents. Some logicians may make extensive use of concrete devices such as whiteboards, sheets of paper and pencils, while others may conduct a significant portion of their thinking without actually writing anything down. Nevertheless, the heuristic function of formal languages is sufficiently pervasive so as to require an explanation. What are the features of formal languages that allow them to perform this operative function?

---

21. Menary analyzes the cognitive impact of writing from the point of view of extended cognition, i.e. the idea that certain external objects, artifacts and technologies are often *constitutive* of the cognitive processes involving them [Menary 2007b].

Before elaborating specifically on the notion of de-semantification, let me point out from the start that this discussion is connected with the 'logic as calculus vs. logic as language' dichotomy introduced by van Heijenoort [Heijenoort 1967]. Typically, the proponents of the use of formal languages in science (logic in particular) emphasize their expressive advantages. Responding to a criticism of the *Begriffsschrift,* Frege makes a point of stressing that his ideography differs from Boole's notation in that it is not a mere calculus: it is a *lingua characteristic* (see [Heijenoort 1967]). It has also often been stressed in the secondary literature that, unlike what was to become of formal languages at a later stage, Frege's logical language is a meaningful language, not a calculus, and that this would be a sign of the superiority of the Fregean project over later projects such as the Hilbertian project (see [Sundholm 2003]). Here, however, I defend a very different view (following [Krämer 2003]): the process of de-semantification of formal languages can be, and often is, *beneficial* for reasoning in and about logic.

That this process of de-semantification did take place historically can be attested by a remark by Tarski (admittedly, with a certain degree of exaggeration), and a famous quote by Whitehead:[22]

> Already at an earlier stage in the development of the deductive method we were, in the construction of a mathematical discipline, supposed to disregard the meanings of all expressions specific to this discipline, and we were to behave as if the places of these expressions were taken by variables void of any independent meaning. But, at least, to the logical concepts we were permitted to ascribe their customary meanings. [...] Now, however, the meanings of all expressions encountered in the given discipline are to be disregarded without exception, and we are supposed to behave in the task of constructing a deductive theory as if its sentences were configurations of signs void of any content. [Tarski 1959, 134]

> By relieving the brain of all unnecessary work, a good notation sets it free to concentrate on more advanced problems, and in effect increases the mental power of the race. In mathematics, granted that we are giving any serious attention to mathematical ideas, the symbolism is invariably an immense simplification. [...] [B]y the aid of symbolism, we can make transitions in reasoning almost mechanically by the eye, which otherwise would call into play the higher faculties of the brain. It is a profoundly erroneous truism [...] that we should cultivate the habit of thinking what we are doing. The precise opposite is the case. Civilization advances

22. In a sense, the very birth of logic as we know it with Aristotle already contained an element of de-semantification, in his extensive use of schematic letters in the two *Analytics* (as pointed out by an anonymous reviewer).

> by extending the number of important operations which we can perform without thinking about them. [Whitehead 1911, 59–61]

But what exactly is meant by 'de-semantification' and what is the connection with the notion of operative writing? Before giving my account of the concept, let me start by offering Krämer's own formulations:

> We can explicate this idea with a type of writing in which this 'process of de-semantification' is particularly apparent. We will name this process 'operative writing'. This modality of writing is commonly known, and misunderstood, as 'formal language' and represents one of the fundamental innovations in 17$^{\text{th}}$ century science. [Krämer 2003, 531]

> The rules of calculus apply exclusively to the syntactic shape of written signs, not to their meaning: thus one can calculate with the sign '0' long before it has been decided if its object of reference, the zero, is a number, in other words, before an interpretation for the numeral '0'—the cardinal number of empty sets—has been found that is mathematically consistent.[23] [Krämer 2003, 532]

> [...] signs can be manipulated without interpretation.[24] This realm separates the knowledge of *how* to solve a problem from the knowledge of *why* this solution functions. [Krämer 2003, 532]

> [...] the potency of these calculations is always connected to a move toward de-semantification: the meanings of signs become un-differentiated. [Krämer 2003, 532]

Essentially, the process of de-semantification entails that each step in the process of reasoning will *not* be guided by the reasoner's intuitions and beliefs which would (almost inevitably, as we shall see) be activated if she was to reason on the basis of the meanings of the inscriptions in question. Rather, by 'mechanically' applying the rules defined within the formalism (as a result of a process of de-semantification; knowing *how* but not necessarily *why*), i.e. by not 'thinking what we are doing', the reasoner not only alleviates the demands on her cognitive resources; she may also effectively block the interference of external information (which, in the case of logical reasoning, should be treated as irrelevant). Moreover, thus defined, operative writing allows for the (in principle) unlimited iteration of such operations; by contrast, if the reasoner is guided by meaning and intuitions, typically at some point she is led too far

23. As reported by [Krämer 2003, fn. 29], until the beginning of modern times in Europe, the zero was not viewed as a number on a par with other numbers, but rather essentially as a 'gap'.

24. Commenting on whether infinitely small or infinitely large numbers are 'actual' numbers in the context of his development of infinitesimal mathematics, Leibniz said: "On n'a point besoin de faire dépendre l'analyse mathématique des controverses métaphysiques" (quoted in [Krämer 2003, fn. 36]).

astray from the cognitive territory she is capable of operating on solely on the basis of her intuitions.

Of course, these points are not entirely new, and the search for the 'mechanization' of reasoning within logic and elsewhere has had its enthusiasts for a long time (even though it has its opponents too, as mentioned above). What I would like to add to the debate in the next section are the findings from psychology and cognitive science that seem to explain, at least partially, why de-semantification and mechanization of reasoning can be such powerful tools, in particular in the quest for *novel* results. But before we move on to psychology, here is one more quote, this time by Gödel:

> [In a formal system] rules of inference are laid down which allow one to pass from the axioms to new formulas and thus to deduce more and more propositions, the outstanding feature of the rules of inference being that they are purely formal, i.e. refer only to the *outward structure* of the formulas, *not to their meanings,* so that they could be applied by someone *who knew nothing* about mathematics, or by a machine. [Gödel 1995, 45; emphasis added]

The theoretical step from disregarding the meaning of formulas and the 'mechanical' application of rules of inference is a substantive step, i.e. de-semantification is neither a necessary nor a sufficient condition for the mechanical application of rules. Nevertheless, the connection between the two notions is a tight one, and it is easy to see that the mechanical application of rules can be greatly facilitated by a process of de-semantification. In the next section we shall see why it can be so beneficial for the production of new knowledge if a given set of rules can in principle be applied for investigations in a given domain by someone who 'knows nothing' about it.

3 Psychology of reasoning and the use of formal languages

In this section, I review some of the results from the psychology of reasoning tradition which, I claim, are relevant for the issues being discussed here, in particular in that they shed light on the operative import of the use of formal languages in logical practices.

Decades of research on the psychology of deduction have amassed significant data on the discrepancies between our spontaneous reasoning mechanisms and the canons of deductive reasoning; the bottom line is that subjects typically do not 'do well' in experiments with deductive tasks (for surveys, see [Evans 2002] and [Dutilh Novaes 2012, chap. 4]). Several explanations for these discrepancies have been put forward, but it seems that we typically let external information 'sneak in', i.e. we take context and prior beliefs into account (even if not directly relevant for the task at hand), just as suggested by Frege in the preface of the *Begriffsschrift.* These are all quite sensible reasoning tendencies for practical purposes, but in scientific contexts they may be counter-productive, for a variety of reasons. In particular, if one of the main goals in science is to produce new knowledge, the tendency we have (see experiments discussed below) to seek confirmation for the beliefs we already hold may hinder scientific discovery. In such contexts, we need devices that help us counter our usual doxastic conservativeness, and formal languages are among such devices.

The psychology of reasoning tradition has also established that the 'mistakes'[25] we make when reasoning (deductively) are not random; there are patterns in such 'mistakes', and to account for the processes which might be underlying this systematicity, the notion of 'reasoning biases' has been introduced. The different biases are presented as an attempt to understand why it is that subjects systematically deviate from normative responses in experiments with deductive tasks. Here I focus on one of these 'biases' recognized in the literature, namely confirmation bias, and more specifically on how confirmation bias affects inference-drawing.

In the abstract of a survey article on confirmation bias, we find the following definition:

> Confirmation bias [...] connotes the seeking or interpreting of evidence in ways that are partial to existing beliefs, expectations, or a hypothesis in hand. [Nickerson 1998, 175]

In practice, people can reinforce their existing beliefs by selectively collecting new evidence, by interpreting evidence in a biased way, by selectively recalling information from memory, etc. In the words of Jonathan Evans, one of the most influential researchers in the field, "confirmation bias is perhaps the best known and most widely accepted notion of inferential error to come out of the literature on human reasoning" [Evans 1989, 41]. 'Confirmation bias' is a general term used to describe different and perhaps unrelated cognitive processes; admittedly, the concept is not always clearly defined. But what the different, possibly unrelated processes collectively referred to as 'confirmation bias' all seem to have in common is what could be described as our overwhelming tendency towards *doxastic conservativeness.* Typically, we seek to confirm and maintain the beliefs we already hold; we are extremely resistant to revising prior beliefs—in C. Fine's [Fine 2006, chap. 4] fitting terminology, our brain is extremely 'pigheaded'.

Among the many guises that confirmation bias can acquire, I am here particularly interested in confirmation bias as affecting inference-drawing. Human agents appear to consistently refrain from drawing conclusions from premises

25. Whether these deviances from the traditional canons of deduction can rightly be considered to be mistakes depends on accepting the normative authority of these canons over thinking and reasoning; but this is a moot point [Elqayam & Evans 2011].

they do accept if the conclusions clash with prior beliefs. They may either draw the 'wrong' conclusions (i.e. those that do not follow from the premises, at least not in the technical sense of deductively following, but instead seem to be in harmony with their prior beliefs) or refuse to draw conclusions in order to avoid conflict with prior belief. The technical term usually used to refer specifically to the phenomenon of prior belief interfering with inference-drawing and inference evaluation (in particular with respect to syllogistic tasks) is belief bias. [26]

The experiments investigating the phenomenon of belief bias are usually formulated with syllogistic arguments; subjects are presented with fully formulated arguments rather than being asked to draw conclusions themselves from given premises. As it turns out, subjects typically endorse far more arguments as valid if their conclusions also accord with prior belief, and reject arguments whose conclusions clash with prior belief, even though they have been instructed to judge only on the basis of what follows 'logically'. [27] Here I report on experiments with evaluation tasks, but belief bias has also been observed in conclusion production tasks, i.e. when subjects are asked to draw conclusions themselves from given premises [Oakhill & Johnson-Laird 1985].

In such studies, subjects are typically presented with syllogisms: some are valid, some are invalid, some have a believable conclusion, some have an unbelievable conclusion, in all four possible combinations. Here are some of the syllogisms presented to subjects in [Evans, Barston, & Pollard 1983] study; notice that the valid syllogisms (both believable and unbelievable) are of the same syllogistic mood and figure (same 'logical form'), and the same holds of the invalid syllogisms.

The general results of the experiment were the following (percentage of arguments accepted as valid):

Clearly, prior beliefs seem to be activated when subjects are evaluating (the correctness of) arguments. Subjects do show a degree of what we could call 'logical competence' (assuming the traditional canons) in that valid syllogisms are more often endorsed than invalid syllogisms in both categories; similarly, syllogisms with believable conclusions are more often endorsed than syllogisms with unbelievable conclusions. What is remarkable though is that *in*valid syllogisms with believable conclusions are more often endorsed (71%) than

26. In fact, the term 'belief bias' is more felicitous as a description of the whole phenomenon than the term 'confirmation bias', given that the bias is not towards confirming *simpliciter* but rather towards reinforcing prior beliefs, which of course may require *dis*confirming (refuting) a particular hypothesis if it clashes with prior beliefs.

27. This is one of the methodological problems with such experiments: they seem to assume that subjects will give an unproblematic interpretation to the notion of 'following logically', but this is arguably a thoroughly theory-laden concept—see [Morris & Sloutsky 1998]. Nevertheless, what emerges from the experiments is still the strong 'pull' of doxastic conservativeness.

Valid-believable	Valid-unbelievable	Invalid-believable	Invalid-unbelievable
No police dogs are vicious.	No nutritional things are inexpensive.	No addictive things are inexpensive.	No millionaires are hard workers.
Some highly trained dogs are vicious.	Some vitamin tablets are inexpensive.	Some cigarettes are inexpensive.	Some rich people are hard workers.
Therefore, some highly trained dogs are not police dogs.	Therefore, some vitamin tablets are not nutritional.	Therefore, some addictive things are not cigarettes.	Therefore, some millionaires are not rich people.

	Believable conclusion	Unbelievable conclusion
Valid	89	56
Invalid	71	10

valid syllogisms with *un*believable conclusions (56%). Apparently, doxastic conservativeness trumps logical competence, and this seems to be an accurate picture of some of our most pervasive cognitive tendencies.[28]

In order to further probe the mechanisms involved in these responses, other researchers [Sá, West, & Stanovich 1999] designed an experiment where syllogisms were presented to subjects with conclusions that would be neither believable nor unbelievable—in other words, syllogisms with so-called 'unfamiliar content', so that the patterns of doxastic conservativeness would presumably not be activated. The prediction was that, if it was indeed doxastic conservativeness that was trumping logical competence, in the case of syllogisms with unfamiliar content this would not occur. In first instance, subjects were presented with an invalid syllogism with familiar content, and in fact with a believable conclusion:

> All living things need water. Roses need water. Roses are living things.

As could have been anticipated, only 32% of the subjects gave the logically 'correct' response when evaluating this syllogism, i.e. that it is invalid. They were then given a little scenario of a situation in a different planet, involving an imaginary species, wampets, and an imaginary class, hudon, and subsequently were asked to evaluate the following syllogism:

> All animals of the hudon class are ferocious. Wampets are ferocious. Wampets are animals of the hudon class.

Interestingly, 78% of the very same subjects whose wide majority had failed to give the 'logically correct' response in the previous task gave the 'logically

28. As described in [Fine 2006, chap. 4], our pigheadedness goes well beyond its manifestation specifically with respect to inference-drawing.

correct' response here, i.e. that the syllogism is invalid. Even more significantly, the two syllogisms have the exact same mood and figure, presumably AAA-2 (the universal quantifiers are unstated in the second premise and the conclusion). So while they had failed to 'see' its invalidity in the first syllogism, arguably in virtue of the familiar content, in the second case the unfamiliar content made it so that prior beliefs did not interfere in the subjects' reasoning. Experiments such as this led K. Stanovich to the following conclusion:

> The rose problem [the discrepancy just mentioned] illustrates one of the fundamental computational biases of human cognition—the tendency to automatically bring prior knowledge to bear when solving problems. That prior knowledge is implicated in performance on this problem even when the person is explicitly told to ignore the real-world believability of the conclusion, illustrates that this tendency toward contextualizing problems with prior knowledge is so ubiquitous that it cannot easily be turned off—hence its characterization here as a fundamental computational bias (one that pervades virtually all thinking whether we like it or not). [Stanovich 2003, 292–293]

(Technically, it is not a matter of knowledge as we philosophers understand it, i.e. as involving factuality, but rather a matter of belief, given that a subject will act in the exact same way towards her false beliefs.) Stanovich then goes on to argue that the computational bias consisting in activating prior beliefs when tackling a given problem is in most cases advantageous, e.g. in terms of the allocation of cognitive resources and time constraints, and that it makes sense to think of it as an adaptation which increased fitness in the human environment of evolutionary adaptedness, many thousands of years ago. But he claims that there are a few (but important) situations in *modern life* where such computational biases must be overridden, as they would provide suboptimal reasoning leading to negative real-life consequences.

It seems to me that, although he may overstate the negative consequences of not overriding 'pigheadedness' in specific situations, there is something very true in Stanovich's observations. As a general cognitive pattern, our tendency towards doxastic conservativeness is normally beneficial; but in specific situations, in particular in scientific contexts, it must be compensated for if we aim at optimal reasoning (given certain goals) in these situations. Which devices could effectively act as a counterbalance to some of these computational biases? My claim here is that formal languages are one such technology, in particular in that the process of de-semantification that is typically (though not always) a core feature of our uses of formal languages as operative devices effectively helps us block the interference of prior beliefs in the reasoning process.

In short, we here see how a philosophical thesis—the claim that Krämer's notions of operative writing and de-semantification can account for the cognitive impact of uses of formal languages (in particular, but not exclusively, in logic)—can receive empirical corroboration. The observed effect of improved logical 'performance' with unfamiliar content in practice seems to correspond to a process of *de-semantification,* and reflecting on doxastic conservativeness as a fundamental computational bias (as defined by Stanovich) gives us a tentative explanation for why our reasoning abilities can be undermined when dealing with familiar content: essentially, because we simply cannot help seeking confirmation for the beliefs we already possess.

Most people would not dispute that the use of formal languages should facilitate the reasoning process in that it augments computational power, as suggested by Whitehead's quote above and as discussed in the literature on the 'extended mind' thesis. But my analysis goes further: the use of formal languages, or 'operative languages' more generally (including mathematical formalisms), in fact allows a human agent to 'run a different software' altogether, as it were, one that is able to override the 'pigheaded' software that we normally run when reasoning. This effect is particularly important in the case of applied logics, i.e. when logic is used to investigate a particular domain or phenomenon, given that in such cases the investigator may have stronger intuitions ('hunches') concerning the target phenomenon in question, which should nevertheless not interfere in the reasoning process. One way to describe the different 'software' induced by reasoning with formal languages is that it forces the agent to consider *all* cases where the premises are true, rather than focusing on her preferred cases (those which accord with her prior, external beliefs).[29]

Similarly, Stenning's [Stenning 2002] results on how the media and the 'languages' used for teaching logic (i.e. sentential or diagrammatical systems of representation) have a significant impact on the students' learning process and reasoning suggest that the 'external means' one uses when doing logic have an undeniable cognitive effect on the reasoning process itself. From this point of view, it is not surprising at all that, once they had the technology of formal languages at their disposal, logicians could do logic in ways that are substantially different from how logic was done prior to the wide adoption of formal languages in logical practices.[30]

---

29. Indeed, one way to conceptualize this difference is in terms of non-monotonicity vs. monotonicity. Our 'computational bias' entails non-monotonic patterns of reasoning, while well-designed formal languages in deductive settings enforce monotonic reasoning—see [Dutilh Novaes 2012] for details.

30. Moreover, work by Landy and Goldstone [Landy & Goldstone 2007a], [Landy & Goldstone 2007b], [Landy & Goldstone 2009] suggests that agents typically call upon sensorimotor processing to operate with mathematical and logical notations, and thus that the perceptual features of the different systems is bound to have an impact on the reasoning processes. They literally 'move bits and pieces' of the notation around to

It might be thought, however, that the results of what is perhaps the most widely studied task in the psychology of reasoning literature, namely the Wason selection task, contradict my main thesis here: subjects' performances on this task are typically very low when they are given 'abstract material' to work with, but they improve in many (though not all) contentful versions of the task.[31]

But what has also emerged from this literature is that not all contentful versions of the task enhance reasoning performance; it is not even the familiarity factor that explains the facilitating effect. What really seems to make a substantive difference is whether subjects interpret the conditional in the task as a descriptive conditional or as a deontic conditional. Stenning and van Lambalgen [Stenning & Lambalgen 2004] have argued that if subjects interpret the conditional as deontic, the task acquires an entirely different logical structure than if they interpret the conditional as descriptive; in fact, with the deontic interpretation, the task is computationally much more tractable—and *not* for reasons associated with confirmation bias or anything related to it. Moreover, the fact that subjects perform poorly when given 'abstract' contents (letters or numbers) only goes on to show that the use of formal languages is not an innate ability in humans: it is a technique that must be learned in order to be used successfully. Thus, the results on the Wason selection task are at the very least compatible with the main hypothesis of this paper: formal languages are a technology which, when properly mastered, may serve as a counterbalance to some of our more spontaneous cognitive patterns, precisely in that they activate other cognitive processes which we *can* perform but which are not activated under normal circumstances.

Conclusion

To conclude, let me refer to a much quoted passage by D'Alembert: 'Algebra is generous: she often gives more than is asked of her'; F. Staal [Staal 2007] has suggested that the same holds of formal/artificial languages. Drawing from empirical data from psychology and cognitive science, I have offered one possible explanation for the generosity of formal languages: they give us more than we expect because they allow us to go beyond the beliefs we already hold; they deliver us new knowledge. This phenomenon is related to the fact

perform the correct transformations. An 'extended cognition' perspective on formal languages is developed in [Dutilh Novaes 2012, chap. 5].

31. For reasons of space, I am assuming that the reader is familiar with the Wason selection task; for further details, see [Stenning & Lambalgen 2004], [Stenning & Lambalgen 2008]. The point of this short discussion is rather to anticipate a natural objection to my main thesis, and to suggest how the objection could be dealt with. To be sure, the discussion of the Wason task here will remain rather superficial for reasons of space; further details can be found in [Dutilh Novaes 2012, chap. 4].

that they are a technology with built-in mechanisms for the inhibition of more spontaneous reasoning tendencies, which seek to confirm prior belief—what I have referred to as our tendency towards doxastic conservativeness.

These mechanisms are related to the operative ('paper-and-pencil') nature of formal languages, i.e. to their role as cognitive artifacts. My suggestion was that these external devices allow us to run a 'different software' when reasoning, a software that is not geared towards bringing in prior beliefs—one of our strongest 'computational biases' according to Stanovich. Notice however that, originally, the avowed purposes of formal languages were essentially expressive and meta-theoretical; in this sense too formal languages are generous, i.e. in the sense of impacting reasoning in ways that go far beyond the original applications they were designed for.

As for the prospects for a practice-based philosophy of logic, I have attempted to illustrate one possible path of development, namely to bring in data about how humans reason in order to attain a better grasp of the analogies and (perhaps more importantly) disanalogies between 'everyday life' forms of reasoning and reasoning as performed by trained logicians. There is every reason to believe that other practice-oriented attempts may deliver significant results, and I hope at least to have convinced the reader that a practice-based approach can be fruitful and shed new light on some long-standing issues within the philosophy of logic—in the present case, why doing logic with formal languages is so different from doing logic without formal languages.

Bibliography

Andrade-Lotero, Edgar & Dutilh-Novaes, Catarina

2010 Validity, the squeezing argument and alternative semantic systems: the case of Aristotelian syllogistic, *Journal of Philosophical Logic*, OnlineFirst, 1–32, doi:DOI10.1007/s10992-010-9166-y.

Avron, Arnon

1994 What is a logical system?, in *What is a logical system?*, edited by Gabbay, D., Oxford: Oxford University Press, 217–238.

Awodey, Steve & Reck, Erich H.

2002 Completeness and Categoricity. Part I: Nineteenth-century Axiomatics to Twentieth-century Metalogic, *History and Philosophy of Logic*, 23, 1–30.

Beall, JC & Restall, Greg

2006 *Logical Pluralism*, Oxford: Oxford University Press.

Benthem, Johan van

2008 Logical dynamics meets logical pluralism?, *Australasian Journal of Logic*, 6, 182–209.

Clark, Andy

2008 *Supersizing the Mind*, Oxford: Oxford University Press.

1998 The extended mind, *Analysis*, 58, 10–23.

De Cruz, Helen, Neth, Hansjörg & Schlimm, Dirk

2010 The cognitive basis of arithmetic, in *PhiMSAMP. Philosophy of mathematics: Sociological aspects and mathematical practice*, edited by Löwe, B. & Müller, T., London: College Publications, 59–106.

De Cruz, Helen & De Smedt, Johan

2010 The innateness hypothesis and mathematical concepts, *Topoi*, 29, 3–13.

Dummett, Michael

1978 The justification of deduction, in *Truth and Other Enigmas*, London: Duckworth, 290–318.

Dutilh Novaes, Catarina

2011 The different ways in which logic is (said to be) formal, *History and Philosophy of Logic*, 32(4), 303–332.

2012 *Formal Languages from a Cognitive Perspective*, Cambridge: Cambridge University Press.

Elqayam, Shira & Evans, Jonathan St.B.T.

2011 Subtracting "ought" from "is": Descriptivism versus normativism in the study of the human thinking, *Behavioral and Brain Sciences*, 34(5), 233–248.

Evans, Jonathan, Barston, Julie & Pollard, Paul

1983 On the conflict between logic and belief in syllogistic reasoning, *Memory and Cognition*, 11, 295–306.

Evans, Jonathan St. B. T.

1989 *Bias in Human Reasoning: Causes and Consequences*, London: Erlbaum Associates.

2002 Logic and human reasoning: An assessment of the deduction paradigm, *Psychological Bulletin*, 128(6), 978–996.

Fine, Cordelia

2006 *A Mind of Its Own: How your brain distorts and deceives*, New York: W.W. Norton.

Frege, Gottlob

1977 Begriffsschrift, in *From Frege to Gödel: A Source Book in Mathematical Logic, 1879-1931*, edited by Heijenoort, Jean van, Cambridge: Harvard University Press, 3rd ed., 1–82, 1879.

Gödel, Kurt
1995 *Collected Works – Vol. III: Unpublished essays and lectures*, Oxford: Oxford University Press.

Haack, Susan
2000 *Philosophy of Logics*, Cambridge: Cambridge University Press.

Heijenoort, Jean van
1967 Logic as calculus and logic as language, *Synthese*, 17, 324–330.

Hersh, Reuben
1997 *What is Mathematics, Really?*, Oxford: Oxford University Press.

Hersh, Reuben (ed.)
2006 *18 Unconventional Essays on the Nature of Mathematics*, Berlin: Springer.

Hodges, Wilfrid
2009 Logic and games, in *Stanford Encyclopedia of Philosophy*, edited by Zalta, E.

Kerkhove, Bart & Bendegem, Jean Paul van
2007 *Perspectives on Mathematical Practices: Bringing together philosophy of mathematics, sociology of mathematics, and mathematics education*, Dordrecht: Springer.

Krämer, Sybille
2003 Writing, notational iconicity, calculus: On writing as a cultural technique, *Modern Languages Notes – German Issue*, 118(3), 518–537.

Kuhn, Thomas
1962 *The Structure of Scientific Revolutions*, Chicago: University of Chicago Press.

Landy, David & Goldstone, Robert L.
2007a How abstract is symbolic thought?, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(4), 720–733.

2007b Formal notations are diagrams: Evidence from a production task, *Memory and Cognition*, 35(8), 2033–2040.

2009 Pushing symbols, in *The 31st Annual Conference of the Cognitive Science Society*, Amsterdam, 1072–1077.
URL `facultystaff.richmond.edu/~dlandy/Publications/Landy_Goldstone_2009.pdf`.

Mancosu, Paolo
2008 Mathematical explanation: Why it matters, in *The Philosophy of Mathematical Practice*, edited by Mancosu, Paolo, Oxford: Oxford University Press, 134–149.

Marion, Mathieu
2009 Why play logical games?, in *Games: Unifying Logic, Language, and Philosophy*, edited by Majer, O., Pietarinen, A.-V., & Tulenheimo, T., Dordrecht: Springer, 3–26.

Menary, Richard
2007b Writing as thinking, *Language Sciences*, 29, 621–632.

Menary, Richard (ed.)
2010 *The Extended Mind*, Cambridge: MIT Press.

Morris, Anne K. & Sloutsky, Vladimir M.
1998 Understanding of logical necessity: Developmental antecedents and cognitive consequences, *Child Development*, 69(3), 721–741.

Mugnai, Massimo
2010 Logic and mathematics in the seventeenth century, *History and Philosophy of Logic*, 31(4), 297–314.

Nickerson, Raymond S.
1998 Confirmation bias: A ubiquitous phenomenon in many guises, *Review of General Psychology*, 2(2), 175–220.

Oakhill, Jane & Johnson-Laird, Philip N.
1985 The effect of belief on the production of syllogistic conclusions, *Quarterly Journal of Experimental Psychology*, 37(A), 553–569.

Popper, Karl
1959 *The Logic of Scientific Discovery*, London: Hutchinson, translation of *Logik der Forschung,* 1934.

Rosental, Claude
2008 *Weaving Self-Evidence: A sociology of logic*, Princeton: Princeton University Press.

Sá Walter West, Richard & Stanovich, Keith
1999 The domain specificity and generality of belief bias: Searching for a generalizable critical thinking skill, *Journal of Educational Psychology*, 91, 497–510.

Schamdt-Besserat, Denise
1996 *How Writing Came About*, Austin: University of Texas Press.

Sher, Gila
2008 Tarski's thesis, in *New Essays on Tarski and Philosophy*, edited by Patterson, D., Oxford: Oxford University Press, 300–339.

Shin, Sun-Joo
2002 *The Iconic Logic of Peirce's Graphs*, Cambridge: MIT Press.

2008 Diagrams, in *Stanford Encyclopedia of Philosophy*, edited by Zalta, E.
URL `plato.stanford.edu/entries/diagrams/`.

Staal, Frits

2006 Artificial languages across sciences and civilizations, *Journal of Indian Philosophy*, 34, 89–141.

2007 Preface: The generosity of formal languages, *Journal of Indian Philosophy*, 35, 405–412.

Stanovich, Keith E.

2003 The fundamental computational biases of human cognition: Heuristics that (sometimes) impair decision making and problem solving, in *The Psychology of Problem Solving*, edited by Davidson, J. E. & Sternberg, R. J., New York: Cambridge University Press, 291–342.

Stenning, Keith

2002 *Seeing Reason*, Oxford: Oxford University Press.

2004 A little logic goes a long way: Basing experiment on semantic theory in the cognitive science of conditional reasoning, *Cognitive Science*, 28, 481–529.

2008 *Human Reasoning and Cognitive Science*, Cambridge: MIT Press.

Sundholm, B. G.

2003 Tarski and Lesńiewski on languages with meaning versus languages without use: A 60th birthday provocation for Jan Woleński, in *Philosophy and Logic. In Search of the Polish Tradition*, edited by Hintikka, J., Czarnecki, T., Kijania-Placek, K., Placek, T., & Rojszczak, A., Dordrecht: Kluwer, 109–128.

Tarski, Alfred

1959 *Introduction to Logic and to the Methodology of Deductive Sciences*, New York: Oxford University Press, 8th ed.

Whitehead, Alfred N.

1911 *An Introduction to Mathematics*, Oxford: Oxford University Press.